图书情报工作
第 64 卷 第 6 期　2020 年 3 月

基于 Word2Vec 和 CNN 的产品评论细粒度情感分析模型*

■ 蔡庆平[1]　马海群[2]
[1] 黑龙江大学信息管理学院　哈尔滨 150080　[2] 黑龙江大学信息资源管理研究中心　哈尔滨 150080

摘　要：[目的/意义]构建一种基于 Word2Vec 和 CNN 的产品评论细粒度情感分析模型。[方法/过程]首先使用 Word2Vec 从产品评论中构建产品特征词列表和噪声词表，其次借助噪声词表来进行产品评论特征词的提取，然后采用 CNN 对产品评论进行产品特征层面的细粒度情感分类，最后实现基于产品特征的产品评论聚类。[结果/结论]通过爬取京东商城华为手机评论对该模型进行训练和测试，结果表明，该模型能够有效实现产品评论的细粒度情感分析，可以有效地发现用户对产品特征的关注度和满意度。
关键词：情感分析　产品评论　CNN　Word2Vec　细粒度
分类号：TP391
DOI：10.13266/j.issn.0252 – 3116.2020.06.007

随着互联网技术在日常生活中的广泛应用，人们越来越习惯于通过网络进行购物，并在购物平台上发表对产品的相关评论，而这些评论通常以文本为载体进行传播。企业通过对产品评论进行情感分析，能够掌握用户对其产品的满意程度，发现产品核心优势和不足之处，并根据用户评论信息适当调整产品战略，这对企业来说是新的机遇和挑战[1]。

1　相关研究与工具选择

1.1　情感分析

情感分析是从评论文本数据中识别出用户所表达的情感倾向的过程。情感分析可以分为粗粒度情感分析和细粒度情感分析两种。

粗粒度情感分析一般是从评论文本整体层面进行情感极性分析，针对整条评论给出积极或者消极的评价。D. T. Vo 等[2]加入表情特征自动构建文本的情感词典来对 Twitter 文本进行情感分析。D. Y. Tang 等[3]通过情感种子扩充特定领域情感词对用户评论进行情感分类。L. Zheng 等[4]利用文档频率选择特征子集，结合支持向量机对中文产品评论进行情感分析。

细粒度情感分析一般是从评论文本片段或者特征层面进行情感极性分析，针对各评论文本片段或者特征分别给出积极或者消极的评价。I. Titov 等[5]构建了一种基于多粒度 LDA 的情感模型（multi-aspect sentiment model，MAS）。孙艳等[6]构建了一种无监督的混合情感模型。Y. Kim[7]将卷积神经网络应用到多个情感分析任务数据集上，并取得了相当优异的效果。李杰等[8]采用卷积神经网络进行短文本评论情感分析可以为产品设计人员进行产品优化改进提供决策支持。

产品评论虽然篇幅较短，但是其中所涉及的产品特征并不唯一，而且用户对各产品特征的态度也不一致，因此产品评论情感分析适合细粒度的情感分析，本文采用卷积神经网络模型实现产品评论的细粒度情感分析。

1.2　产品评论特征词提取

产品评论特征词提取是指从产品评论文本中提取用户评价的、与产品特征相关的词语，通常包括产品的功能、性能等。针对产品评论特征词提取，很多学者进行了相关的研究。余传明等[9]基于支持向量机从客户评论文本中提取产品特征词。徐建民等[10]在本体库的基础上，利用 TF-IDF 方法实现了对文本中特征词的提取。夏天[11]将词向量技术融入 TextRank 方法，提高了特征词提取的效果。

随着深度学习在情感分析领域的广泛应用，基于

* 本文系国家社会科学基金重点项目“开放数据与数据安全的政策协同研究”（项目编号：15ATQ008）和黑龙江省省属高等学校基本科研业务费基础研究项目“基于深度学习的产品评论信息情感分析研究”（项目编号：RWSKCX201809）研究成果之一。
作者简介：蔡庆平（ORCID：0000 – 0002 – 1992 – 4081），讲师，博士研究生，E-mail：cqp-cqp@163.com；马海群（ORCID：0000 – 0002 – 2091 – 7620），教授，博士，博士生导师。
收稿日期：2019 – 05 – 19　修回日期：2019 – 12 – 02　本文起止页码：49 – 58　本文责任编辑：易飞

概率的浅层神经网络模型悄然兴起[12]，Word2Vec 是这类模型的佼佼者，越来越多的学者研究使用 Word2Vec 进行文本特征词的提取。李跃鹏等[13]使用 Word2Vec 技术结合 kmeans 聚类方法提出一种特征词提取算法。周顺先等[14]使用 Word2Vec 构建词向量聚类质心频率模型，用于文本的特征词提取。

本文针对产品评论进行细粒度情感分析，产品评论特征词提取是细粒度情感分析的重要基础，因此产品评论特征词提取的效果尤为重要，本文尝试使用 Word2Vec 来进行产品评论特征词提取，着力提高产品评论特征词提取的效果。

1.3 CNN（卷积神经网络）

本文选择 CNN（convolution neural network，卷积神经网络）[15-16]作为产品评论的情感分类器。图 1 为用于产品评论文本分类的 CNN 模型结构，CNN 模型由输入层、卷积层、池化层和全连接层组成。

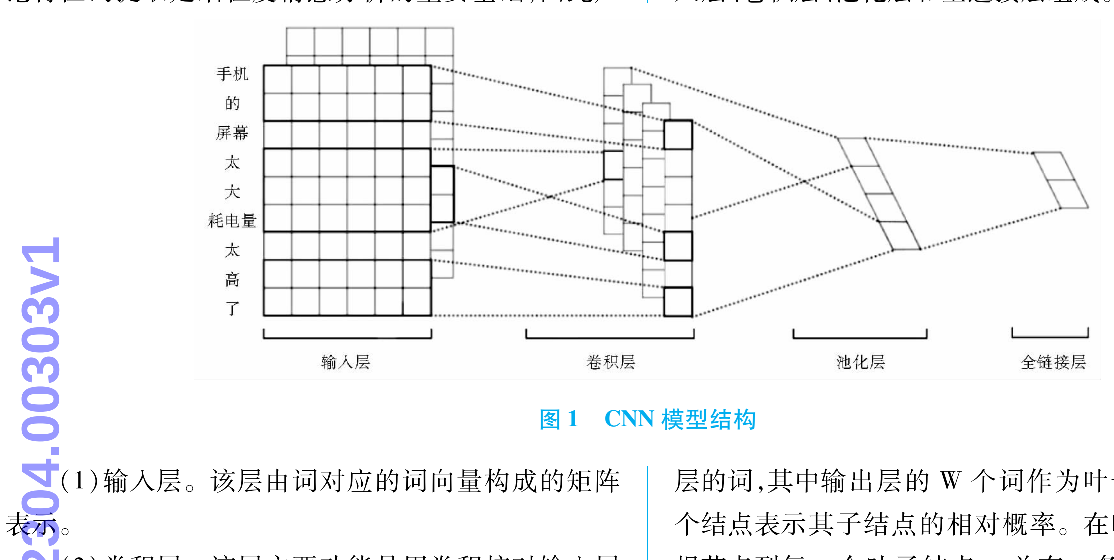


图 1 CNN 模型结构

（1）输入层。该层由词对应的词向量构成的矩阵表示。

（2）卷积层。该层主要功能是用卷积核对输入层词向量矩阵进行卷积操作，得到更深层次的文本特征。

（3）池化层。该层的主要功能是对卷积层提取到的文本向量进行特征选择和信息过滤，在保留主要特征的同时减少下一层的参数和计算量，防止过拟合。

（4）全连接层。该层主要功能是将从池化层得到的 m 个产品评论中最显著的文本特征所对应的特征值进行全连接操作，以得到一个固定长度的特征向量来表示产品评论的特征，将特征向量输入到最后的 softmax Regression 分类器中，从全局的角度对特征进行分析，进而完成产品特征评论的情感分类。

1.4 Word2Vec

Word2Vec[17-20]是 Google 2013 年提出的一种词语语义计算技术。通过 Word2Vec 训练，可以把文本内容的处理简化为 K 维向量空间中的向量运算，向量空间上的相似度可以用来表示文本语义上的相似度。Word2Vec 提供了两种经典的语言模型进行训练：CBOW 模型和 Skip-gram 模型。针对这两个模型，Word2Vec 给出了两个框架，分别是基于 Hierarchical Softmax 和 Negetive Sampling 来进行设计，本文采用基于 Hierarchical Softmax 的 Skip-gram 模型。

Hierarchical Softmax 使用哈夫曼树结构表示输出层的词，其中输出层的 W 个词作为叶子结点存在，每个结点表示其子结点的相对概率。在哈夫曼树中，从根节点到每一个叶子结点 w 总有一条最适当的路径存在。

Skip-gram 模型由三层网络模型构成，即输入层、投影层、输出层，如图 2 所示。Skip-gram 模型的训练目标是寻找到有助于预测句子或文档中周围单词的单词表示。

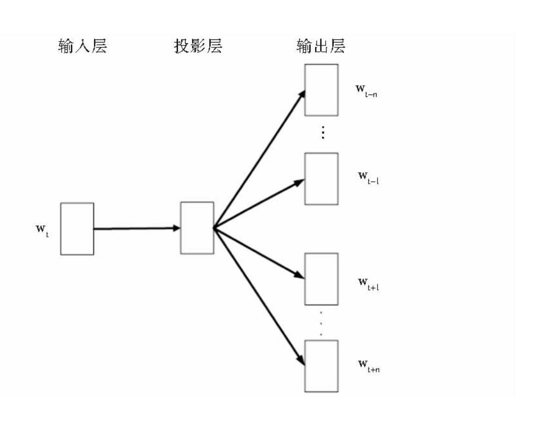


图 2 Skip-gram 模型结构

2 基于 Word2Vec 和 CNN 的产品评论细粒度情感分析模型

本文设计了一个基于 Word2Vec 和 CNN 的产品评论细粒度情感分析模型，包括产品评论预处理模块、产

蔡庆平，马海群. 基于 Word2Vec 和 CNN 的产品评论细粒度情感分析模型[J]. 图书情报工作,2020,64(6):49－58.

品特征词表和噪声词表构建模块、产品评论特征词提取模块、产品评论情感分类模块、基于产品特征的产品评论聚类模块 5 个部分，详如图 3 所示：

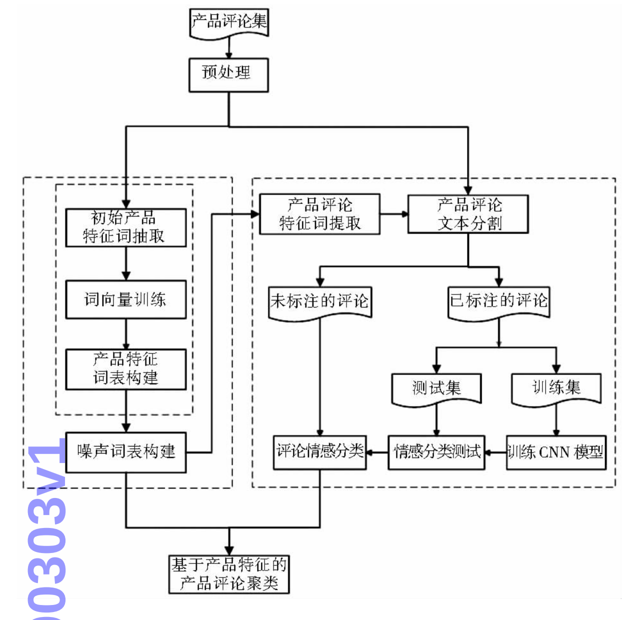


图 3　基于 Word2Vec 和 CNN 的产品评论细粒度情感分析模型

2.1　产品评论预处理

产品评论集中往往存在一些无意义的产品评论，这些评论中不包含任何特征词，比如只包含“好”“差评”“不错”“点个赞”等词，因此过滤掉这些无用的产品评论有利于减少噪声干扰、提高文本情感分类和特征词提取的准确率。

由于中文不像英文以空格作为单词之间的分隔符，中文词汇之间没有明确的界限，因此需要先对产品评论进行中文分词和词性标注，以词作为产品评论的组成要素。

产品评论分词和词性标注完毕后，还要借助停用词表去除产品评论中的常见停用词，常见停用词包括“突然”“立刻”“不但”“而且”“我们”等与产品特征、情感分类无关的一些词。

2.2　产品特征词表和噪声词表构建

2.2.1　产品特征词表构建

产品特征是指产品所具备的属性或功能，以手机为例，屏幕、外观、内存、摄像头等均属于手机的产品特征。产品特征词表是构建噪声词表和基于产品特征的产品评论聚类的依据。产品特征词表构建分为 3 个部分：

（1）初始产品特征词抽取。由于多数产品的属性词为名词和名词短语[21]，而高频名词和名词短语往往是真正的特征词，因此本文对预处理后的产品评论中的名词和名词短语进行词频统计，人工从中抽取前 m 个与产品特征相关的高频名词和名词短语作特征词，将这些特征词按照产品特征的类别进行分类，构成初始产品特征词表 Ts。

（2）词向量训练。词间相似度量可以方便地通过计算词向量的余弦距离来度量，因此词向量表示了语料中词与词间的深层语义联系。Word2Vec 是以由文本数据构建的词汇表为训练数据，然后学习词的高维向量表示，即将词映射至有限维的高维空间中。Word2Vec 工具对进行分词处理后的产品评论集进行训练，可以获得词向量模型以及每个词的指定维度的向量表示。词向量既可以用于产品特征的提取，也可用于卷积神经网络的输入。

（3）产品特征词表的生成。描述同一产品特征的特征词可能不是一个，例如描述手机外观特征的特征词有外观、外观设计、造型、颜值、款式等，因此要想获取较为全面的特征词，就需要对特征词进行聚类。本文使用 Word2Vec 训练获得的词向量模型计算产品评论集中不同词与初始产品特征词表 T_s 中的每个特征词之间的相似度，选择与每个产品特征词相似度较高的前 n 个与产品特征相关的词来扩充初始产品特征词表 T_s，进而完成产品特征词表 T 的生成。

2.2.2　噪声词表构建

由于从产品评论中提取的特征词中可能不全是特征词，其中包含一部分跟产品特征无关的词，本文认为这部分词是噪声词，会对基于产品特征的产品评论聚类过程产生干扰，所以本文通过构造噪声词表的方式来过滤从产品评论文本中提取的特征词序列，去除其中的噪声词进而正确地提取产品评论特征词。噪声词表构建过程如下：

输入：产品特征词表 T 和产品评论集 D 经过 Word2Vec 训练后得到的词向量模型

Step1：从 T 中读取任意特征词 w_i。

Step2：计算与 w_i 相似度较高的前 n 个词。

Step3：从这 n 个词中寻找和产品特征无关的词加入到噪声词表 Z 中。

Step4：重复 Step1、Step2 和 Step3，直到 T 中的每一个产品特征词都处理完成。

输出：噪声词表 Z。

2.3　产品评论特征词提取

本文中产品评论特征词提取方法基于 Hierarchical Softmax 的 Skip-gram 模型实现，其过程如下：

输入：预处理后的产品评论集 $D=\{S_1,S_2,\cdots,S_m\}$。

Step1：读取 D 中的任意一条产品评论 S_k，其中（$k=1,2,\cdots,m$），$S_k=\{w_1,w_2,\cdots,w_i,\cdots,w_{nk}\}$，针对 S_k 中的每个词 w_i，计算其 $p(w_j|w_i)$，其中 $j=1,2,\cdots,nk$。

Step2：针对 S_k 中的每个词 w_i，计算 $p(S_k|w_i)$，其中

$$p(S_k|w_i)=p(w_1,w_2,\cdots,w_{nk}|w_i)=\prod_{j=1}^{nk}p(w_j|w_i) \tag{1}$$

Step3：将 S_k 中所有词的 $p(S_k|w_i)$ 降序排列，排好序的前 n 个词选择作为 S_k 的特征词。

Step4：重复 Step1、Step2 和 Step3，直到 D 中的所有产品评论都完成特征词提取，形成产品评论特征词列表 D_{key}。

Step5：使用噪声词表 Z 过滤产品评论集 D 所对应产品评论特征词列表 D_{key}，最终获得去除噪声的产品评论特征词列表 D_c。

输出：产品评论集 D 所对应产品评论特征词列表 D_c。

2.4 产品评论情感分类

2.4.1 产品评论文本的分割

每一条产品评论中包含的产品特征可能不是一个，每个产品特征的情感标签也可能是不同的。比如“屏幕足够大，流畅性不错，信号不错，但是电池不耐用”，这条产品评论中包含“屏幕”“性能”“信号”“电池”等特征，其中“屏幕、性能、信号”特征的情感标签是积极的，“电池”特征的情感标签是消极的。

一条产品评论文本中往往包含多个分隔符，这些分隔符包括：（。）、（，）、（！）、（#）、（*）以及（常见表情符）等，而用户表达对某一产品特征观点的短句通常在一个分隔符之内，所以本文采用基于分隔符分割的策略（BSP）[22]，将每一条产品评论分割成多个短句，然后从多个短句中过滤掉不包含产品特征词的短句，如果有多条短句描述同一产品特征则将其合并成一个新的短句，最终使每一条短句只评价一种产品特征。产品评论集 D 中的评论文本经过分割后形成新的产品评论集 D_w。

2.4.2 CNN 模型训练和测试

从产品评论集 D_w 中提取部分产品评论数据作为训练集，进行 CNN 模型的情感分类训练。训练集中的样本已通过人工标注形式添加了情感标签，使用 CNN 模型计算得到训练样本的情感标签，并与已有情感标签比较，利用计算误差不断对模型参数进行调整。从产品评论集 D_w 中提取另一部分产品评论数据作为测试集，分类训练完成后，使用测试集对 CNN 模型的情感分类效果进行评价，即利用 CNN 模型计算得到测试集中产品评论的情感标签，并与其人工标注的情感标签比较。

2.4.3 产品评论情感分类

CNN 模型完成训练和测试以后，可以将产品评论集 D_w 中未做情感标注的评论文本进行情感分类，为每一条产品评论添加情感标签。

2.5 基于产品特征的产品评论聚类

经过 CNN 模型的情感分类之后，产品评论集 D_w 中的每一条产品评论都带有相应的情感标签，其中的特征词也带有相应的情感标签。为了确定每条产品评论对哪种产品特征进行了积极还是消极的情感评价，本文首先对产品评论集 D_w 中的每一条产品评论进行特征词提取，然后将产品评论进行基于产品特征的聚类，以用来分析和评价用户对产品特征的关注度和满意度。基于产品特征的产品评论聚类过程如下：

输入：产品评论集 D_w 和产品特征词表 $T=\{T_1,T_2,\cdots T_q\}$，其中 T_j 代表第 j 个产品特征所对应的特征词列表。

Step1：提取 D_w 中每一条产品评论的特征词，生成产品评论特征词列表 $D_{split}=\{S_1,S_2,\cdots,S_p\}$。

Step2：读取 D_{split} 中的任意一行 S_i，$S_i=\{w_1,w_2,\cdots,w_{ni}\}$，其中（$i=1,2,\cdots,p$）。

Step3：读取 S_i 中的任意特征词 w_i。

Step4：读取 T 中任何一个 T_j，$T_j=\{t_1,t_2,\cdots,t_{nj}\}$，其中（$j=1,2,\cdots,q$）。

Step5：读取 T_j 中的任意特征词 t_o。

Step6：计算 w_i 与 t_o 之间的相似度 Sim_i^o。

Step7：重复 Step5、Step6，记录其中的最大值 $\max_{ij}=\max(\{Sim_i^1,Sim_i^2,\cdots Sim_i^{nj}\})$。

Step8：重复 Step4 到 Step7，计算 $\max(\{\max_{i1},\max_{i2},\cdots,\max_{iq}\})$，可以确定特征词 w_i 所属的特征类别。

Step9：重复 Step3 到 Step8，直到 S_i 中所有的词都读取完成，确定评论文本 S_i 归属的产品特征类别。

Step10：重复 Step2 到 Step9，直到 D_{split} 中所有行处理完成。

输出：输出各产品评论所属产品特征类别。

3 实验与结果分析

3.1 实验环境

本文所有实验均在表 1 环境中完成。

蔡庆平，马海群. 基于 Word2Vec 和 CNN 的产品评论细粒度情感分析模型[J]. 图书情报工作,2020,64(6):49－58.

3.2　产品评论预处理

本文通过 Python 语言编写网络爬虫程序采集了京东商城华为手机产品评论共 10 223 条，本文将以上述华为手机评论集为例进行模型应用和结果分析。

首先，对一些无意义的产品评论进行清洗，最终过滤后的得到产品评论集共 9 230 条。

表 1　实验环境与配置

实验环境	配置参数
操作系统	Win7
CPU	Intel(R) Core(TM) i3-4130 3.40GHz
内存	8G
深度学习框架	Tensorflow 1.3-cpu
程序设计语言	Python 3.6
分词工具	jieba
词向量训练工具	Word2Vec(gensim2.3.0)
编程环境	Anaconda

其次，文中实验采用 Python 语言的 jieba 分词包对产品评论集中的评论文本进行分词并标注词性。

最后，本文以哈尔滨工业大学停用词表为基础，通过对分词后的产品评论集中的词进行词频统计，选择高词频中与产品特征和情感词无关的词来构建适合手机评论的停用词表，并使用该停用词表对分词后的产品评论集进行去停用词处理。

3.3　产品特征词表和噪声词表构建

3.3.1　词向量训练

本文利用进行分词处理后的产品评论集作为词向量训练语料，使用 Python 语言的 gensim 包中的 Word2Vec 库中自带的函数来训练词向量，构建词向量模型。表 2 为 Word2Vec 参数设置，表 3 中显示的是产品评论集中部分词的词向量：

表 2　Word2Vec 参数设置

参数	参数值
算法选择(sg)	Skip-gram
词向量维度(size)	128
词频最小值(min_count)	3
训练窗口大小(window)	5
并行线程数(workers)	当前运行机器的处理器核数

表 3　词向量列表(部分)

词	词向量				
屏幕	－0.125 163 87	－0.024 804 92	0.129 791 13	0.123 645 24	0.212 407 13
速度	－0.118 312 66	0.051 090 38	－0.001 206 69	0.009 693 74	0.180 293 86
电池	－0.141 637 13	0.163 745 14	0.049 230 19	0.212 385 18	0.162 825 32
外观	－0.105 057 77	0.002 056 10	－0.021 922 46	－0.058 548 31	0.141 570 78
性价比	－0.086 082 29	0.03 331 02	0.039 414 20	－0.023 441 83	0.141 992 39
摄像头	0.016 293 62	－0.004 493 70	0.239 884 7	－0.118 341 22	－0.146 049 7
像素	0.048 881 55	0.107 083 5	0.056 056 09	－0.048 168 75	－0.102 540 2

3.3.2　产品特征词表和噪声词表的生成

首先对预处理后的产品评论集中的名词和名词短语进行词频统计，并抽取其中前 200 个与手机特征相关的高频名词和名词短语作为初始产品特征词，将这些特征词按照产品特征的类别进行分类，形成初始产品特征词表，详见表 4；其次使用 Python 语言的 gensim 包中的 Word2Vec 所自带的函数 most_similar( ) 来计算产品评论集中的名词和名词短语与初始产品特征词之间的相似度，在与产品特征词相似度较高的前 50 个词中选择与产品特征真正相关的词来扩充初始产品特征词表，图 4 显示的是与特征词“网络”相似度较高的前 50 个词的计算过程，结果见表 5；然后从表 5 中选择与“网络”相关的词扩充到初始产品特征词表中，生成产品特征词表，详见表 6；最后在产品特征词表基础上构建噪声词表，详见表 7。

表 6 中的每一行表示一种产品特征及其所对应的特征词列表，表 7 中的每一行表示一种产品特征及其相关的噪声词列表，产品评论共涉及屏幕、性能、网络、摄像头、电池、外观、功能、通话、质量、配件 10 类特征。

表 4　初始产品特征词表(部分)

产品特征	初始产品特征词
屏幕	屏幕\分辨率\触屏\色彩\画面\全屏\屏幕显示\斜纹\清晰度
电池	电池\电量\充电器\待机\待机时间\手机电池\小时
网络	信号\网络\网页\网速

```
import numpy as np
import gensim
model = gensim.models.word2vec.Word2Vec.load('./data1/word2vec/data.model')
y2= model.most_similar(u"网络",topn=50)
print("和【网络】最相关的词有：\n")
print("候选特征词        相似度")
for item in y2:
    print("%s          %f'  %(item[0],item[1]))
```

图 4　Word2Vec 产品特征词聚类

表 5 "网络"特征相关的候选特征词

候选词	相似度	候选词	相似度	候选词	相似度
信号	0.954 357	技术员	0.768 731	稳定性	0.739 962
手机信号	0.942 106	信息	0.765 410	重影	0.737 712
电信卡	0.904 950	接收器	0.763 482	网页	0.737 200
满格	0.898 380	热点	0.761 770	个差	0.735 901
路线	0.880 930	排行榜	0.759 959	信用	0.734 552
通话质量	0.865 511	信号强度	0.758 529	音色	0.733 752
参数	0.835 826	流量	0.757 990	平板	0.732 208
隧道	0.833 678	工程	0.751 472	单卡	0.723 977
断网	0.822 136	一格	0.751 411	界面显示	0.717 993
断流	0.806 004	手环	0.748 245	运营商	0.716 717
无线	0.793 062	主卡	0.747 787	数据	0.716 339
无线网络	0.792 115	邮件	0.745 350	熄屏	0.712 905
无线网	0.784 521	面子	0.745 085	视频	0.712 050
温度	0.779 664	重度	0.744 010	网速	0.711 658
网络连接	0.779 486	插件	0.743 536	框框	0.710 977
链接	0.770 942	下文	0.742 024	耳麦	0.710 408
后台程序	0.768 898	讲话	0.740 15		

表 6 产品特征词表(部分)

产品特征	产品特征词
屏幕	屏幕\分辨率\触屏\色彩\画面\全屏\屏幕显示\液晶\饱和度\网格\拖影\视觉效果\阴影\显示屏\斜纹\霸屏\密度\色温\颗粒\清晰度\灵敏度\色差\主屏\内屏 外屏\前屏\色调
电池	电池\小时\电量\充电器\待机\待机时间\手机电池\电池容量\用电量\温度\耗电量\费电\用电\低电量\发热量
网络	信号\手机信号\断网\无线\无线网络\无线网\网络连接\链接\热点\信号强度\流量\网页\网速\接收器

表 7 噪声词表(部分)

产品特征	噪声词
屏幕	英寸\宽度\视野\模组\物体\文件夹\风格\光线
电池	能力\强度\电子书\游戏\大陆\电玩
网络	路线\隧道\断流\主卡\电梯\运营商

3.4 产品评论特征词提取

(1)借助于已经训练好的词向量模型,提取每一条产品评论中所包含的产品评论特征词,形成产品评论集所对应的产品评论特征词列表,如表 8 所示:

表 8 产品评论特征词列表(部分)

产品评论	产品评论特征词
评价晚了。京东物流真的很给力,上午快 11 点下单,下午就送到了。头一次尝试荣耀系列,宝贝不错,屏幕做的很给力,分辨率也很好高,这个价位有这样的机子,真心不错,用麒麟的处理器速度也很快。网评这款处理器听挺不错的。后期看看用起来会不会卡,好用就以后可以支持华为系列啦;	麒麟\处理器\速度\分辨率\网评\屏幕\机子
手机外观漂亮,黑色版本的,屏幕色彩也很鲜艳,屏幕大而机身轻薄,手机运行速度也比较快,很对得起这个价格,值得推荐,京东物流也特别快,赞一个;	黑色\外观\速度\屏幕\色彩\版本\机身
无线信号太弱,无法忍受,坐在同一位置手机离远与离近竟然相差两格,路由器隔墙的情况下。这与发布会上说的高铁不掉线也差距太了吧?	信号\无线\路由器

(2)使用噪声词表过滤产品评论特征词列表中的噪声词,生成去除噪声的产品评论特征词列表,如表 9 所示:

表 9 去除噪声的产品评论特征词列表(部分)

产品评论	产品评论特征词
评价晚了。京东物流真的很给力,上午快 11 点下单,下午就送到了。头一次尝试荣耀系列,宝贝不错,屏幕做的很给力,分辨率也很好高,这个价位有这样的机子,真心不错,用麒麟的处理器速度也很快。网评这款处理器挺不错的;	麒麟\处理器\速度\分辨率\屏幕
手机外观漂亮,黑色版本的,屏幕色彩也很鲜艳,屏幕大而机身轻薄,手机运行速度也比较快,很对得起这个价格,值得推荐,京东物流也特别快,赞一个;	黑色\外观\速度\屏幕\色彩
无线信号太弱,无法忍受,坐在同一位置手机离远与离近竟然相差两格,路由器隔墙的情况下。这与发布会上说的高铁不掉线差距也太大了吧?	信号\无线

为了便于与已有的特征词提取方法对比分析,本文采用精准率($P_{extract}$)、召回率($R_{extract}$)和 F1 值($F1_{extract}$)作为特征词提取效果的评估标准,$P_{extract}$、$R_{extract}$ 和 $F1_{extract}$ 的计算方法如公式(2)所示[11,23],其中 c_i 表示第 i 条产品评论通过特征提取方法提取出来的特征集合,d_i 表示第 i 条产品评论自身所附带的特征词集合,M 表示待处理产品评论集中产品评论数量。

$$P_{extract}=\frac{\sum_{i=1}^{M}\frac{|c_i\cap d_i|}{|c_i|}}{M},\ R_{extract}=\frac{\sum_{i=1}^{M}\frac{|c_i\cap d_i|}{|d_i|}}{M},\ F1_{extract}=\frac{2\times P\times R}{P+R} \quad (2)$$

从产品评论集中选取 500 条产品评论,分别使用 TF-IDF 方法、TextRank 方法和本文的特征词提取方法进行处理,3 种方法设定特征词提取的数量为 2 - 10 个,分别计算 $P_{extract}$、$R_{extract}$ 和 $F1_{extract}$ 3 个评估指标,然后进行对比,最终的结果见图 5、图 6 和图 7。

从图 5、图 6 和图 7 可以看出,在使用本文的特征词提取方法进行特征词提取时,分别设置不同的特征

词提取数量，其精准率（$P_{extract}$）和 *F1* 值（$F1_{extract}$）均优于 *TF-IDF* 方法、*TextRank* 方法，召回率（$R_{extract}$）在特征词数量为 8－10 时也与 *TF-IDF* 方法、*TextRank* 方法持平。所以针对产品评论这类短文本，使用本文的基于噪声词表过滤的特征提取方法进行特征提取效果较为显著。

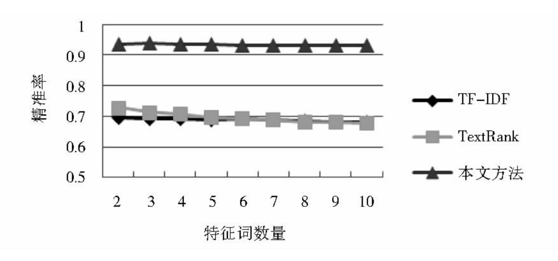


图 5　精准率 $P_{extract}$

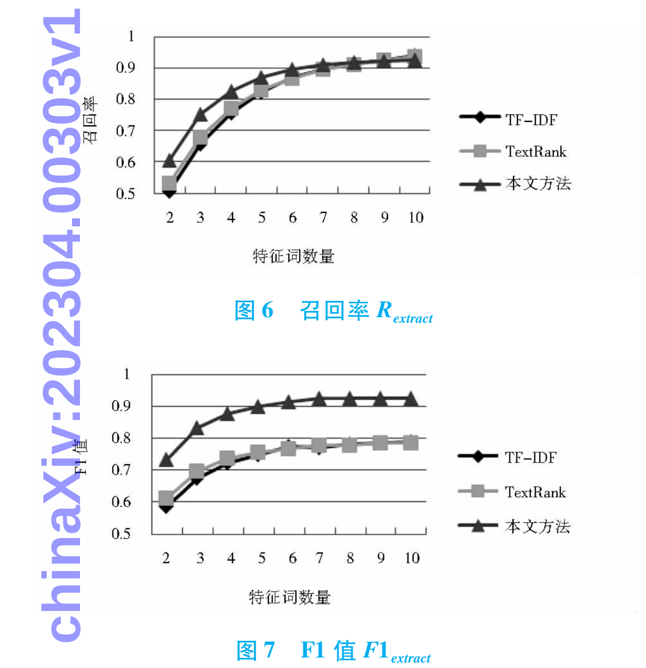


图 6　召回率 $R_{extract}$

图 7　F1 值 $F1_{extract}$

3.5　产品评论情感分类

CNN 模型开始训练前，需要确定其参数。CNN 模型主要具有以下参数：①卷积核尺寸，即对输入词向量进行卷积的区域大小。词语作为最小语言单位不做划分，只需要考虑语境上下文的影响，因此卷积核宽度保持最大不变，只改变卷积核高度。②卷积核数量。③dropout 比例，即取值置为 0 的数据比例。④随机梯度下降算法的批量值。本文采用网格搜索法[24]确定以上参数，而随机梯度下降算法的学习率使用 Adadelta 更新规则进行自适应调整，最终确定的参数取值见表 10。

表 10　卷积神经网络（CNN）参数设置

参数	参数值
卷积核尺寸（kernel_size）	（3,128），（4,128），（5,128）
卷积核数量	256
dropout 比例	0.5
批量值（batch_size）	64

首先将产品评论集中的每一条产品评论进行分割，分割结果详见表 11，然后将分割后的产品评论集进行预处理，并按照 6:2:2 的比例划分为训练集、验证集和测试集 3 部分。训练集作用是用来拟合模型，通过设置分类器的参数，训练分类模型；验证集的作用是使用训练出来的模型对验证集数据进行预测，用来调整模型参数，选出效果最佳的模型所对应的参数；测试集作用是使用训练好的 CNN 模型为未做情感分类的产品评论进行情感分类。

情感分类实验结果评价指标选择信息检索领域传统的精准率（*Precision*）、召回率（*Recall*）、F1 值（F1-score）[15]。设置 SVM 和 NB（Naïve Bayes）两个对照实验组，与 CNN 进行对照，计算 *Precision*、*Recall* 和 *F1-score* 指标，比较三种分类器的性能。CNN 情感分类结果和对照实验的对比结果详见表 12。从表 12 可以看出，在 *Precision*、*Recall* 和 *F1-score* 方面，CNN 对产品评论文本进行情感分类的效果更好。

表 11　产品评论分割（部分）

产品评论	分割后的产品评论
评价晚了。京东物流真的很给力，上午快 11 点下单，下午就送到了。头一次尝试荣耀系列，宝贝不错，屏幕做的很给力，分辨率也很好高，这个价位有这样的机子，真心不错，用麒麟的处理器速度也很快。网评这款处理器挺不错的；	屏幕做的很给力 分辨率也很好高 用麒麟的处理器速度也很快，网评这款处理器挺不错的
手机外观漂亮，黑色版本的，屏幕色彩也很鲜艳，屏幕大而机身轻薄，手机运行速度也比较快，很对得起这个价格，值得推荐，京东物流也特别快，赞一个；	手机外观漂亮，黑色版本的 屏幕色彩也很鲜艳，屏幕大而机身轻薄 手机运行速度也比较快
无线信号太弱，无法忍受，坐在同一位置手机离远与离近竟然相差两格，路由器隔墙的情况下。这与发布会上说的高铁不掉线差距也太大了吧？	无线信号太弱

3.6 基于产品特征的评论聚类

产品评论集中的评论文本经过分割后会生成产品评论短句集，这些短句经过 CNN 模型情感分类后，每一条短句中描述的产品特征就带有相应的情感标签。

表 12 情感分类实验对比结果

分类模型	准确率(P)	召回率(R)	F1 值
NB	82.52%	82.46%	82.47%
SVM	83.43%	83.29%	83.28%
CNN	86.27%	88.39%	87.27%

使用 Python 语言的 gensim 包中 Word2Vec 自带的 similarity( )函数计算每条短句中包含的特征词与表 6 中各类产品特征的特征词之间的相似度，选择其中相似度最大的作为当前短句的归属特征类别，进而将各产品评论基于产品特征进行聚类，并进行相关统计分析，统计结果详见如图 8 和图 9 所示：

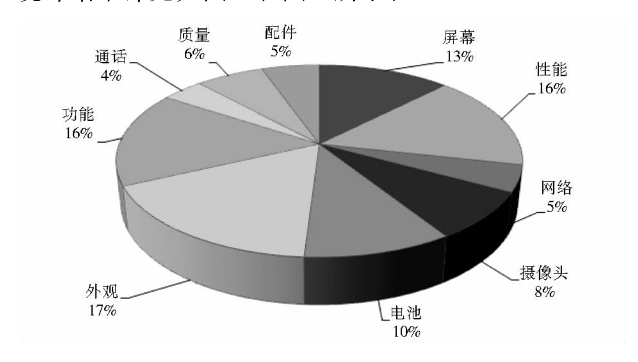


图 8 手机特征评论分布

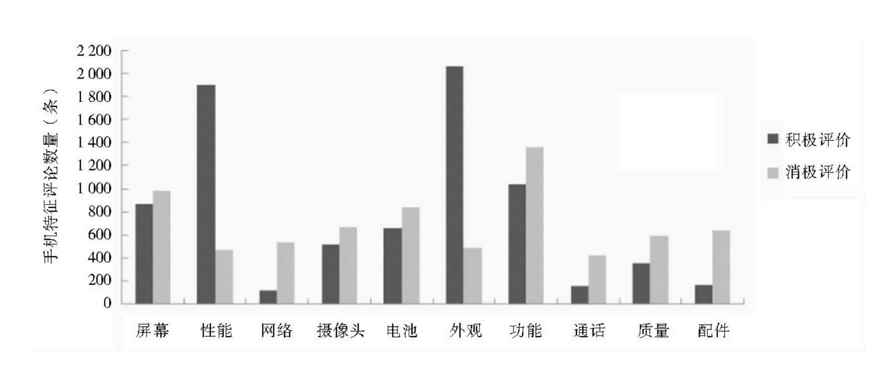


图 9 华为手机特征评论情感倾向

3.7 实验结果分析

图 8 和图 9 展示了用户对华为手机特征的相关评论的分布情况以及情感倾向。从图 8 可以看出，用户重点关注的产品特征是屏幕、性能、外观、功能。从图 9 可以看出，用户对手机性能和外观的满意度较高，在网络、通话和配件方面用户诟病很多，屏幕、摄像头、电池方面用户评论好坏参半。

在性能方面，华为自主研发的麒麟系列处理器性能接近高通骁龙系列处理器，同价位的华为手机配置相对要高于其他品牌手机，因此手机性能较为卓越；在外观方面，华为的荣耀系列采用炫彩的玻璃机身设计，时尚、漂亮，符合年轻人的审美需求，深受年轻人的喜爱。对于此类产品特征，设计人员可以保持设计方案。

华为莱卡摄像头的出现很符合用户需求，但是拍照反应迟钝、夜间成像的效果有待改进；用户对手机屏幕的校准、指纹识别、灵敏度不高和易碎等问题关注较多；在功能方面，缺少双击锁屏、NFC、指纹支付等功能；在电池方面，用户普遍反映电池容量 4000mAh，但没有想象的那么好，好像 3200mAh，续航较差、充电发热较为严重。设计人员应在保持优点的同时尽快对缺点进行改进。

对于网络、通话和配件这种消极评价较多的产品特征，设计人员应迅速寻原因加以改进。例如，用户反映手机接收 Wifi 信号能力差、切换 4G 网络也经常出现卡顿、玩游戏经常掉线等问题，由于华为主营业务为通讯，用户对手机网络这一产品特征期望会比较高；在通话方面，用户更多地反映有噪声、音质不清晰、音量小等问题；另一方面，按键、耳机、听筒、保护膜等配件较为容易损坏。设计人员可以通过改进网络、通话和配件方面的问题来进一步提升手机的竞争力。

在网络盛行的今天，网络平台上充斥着大量的产品评论，产品评论的细粒度情感分析对用户和企业来说显得尤为重要。一方面，用户通过对产品评论进行细粒度情感分析，可以了解到其他用户的使用经验以及对产品相关特征的评价，更有利于用户在购买产品时做出正确的决策；另一方面，产品的口碑对企业来说至关重要，正面的口碑有助于企业产品的推广和销售，负面的口碑则会损害企业的形象。因此企业通过对产品评论进行细粒度情感分析，可以动态掌握用户对产品的口碑变化，当负面口碑出现时，能够及时预警并做

好应对措施,减少对企业的影响。

4 结语

本文对在线产品评论情感分析问题进行了研究,提出了一种基于 Word2Vec 和 CNN 的产品评论细粒度情感分析模型。模型实现了产品评论的特征词提取、产品评论特征层面的情感分类,并在此基础上实现了基于产品特征的产品评论聚类分析。该模型能够实现产品评论的细粒度情感分析,比较全面地捕获产品评论中的产品特征的情感信息,可以较高效地从产品评论中挖掘出用户对产品特征的关注度和满意度,为企业管理和决策提供服务。

本文的创新点在于,针对产品评论这类短文本的特点,提出一种使用 Word2Vec 技术的基于噪声词过滤的产品评论特征词提取方法,该方法不仅充分考虑到了上下文语义信息,而且进一步提高了产品特征词提取的效果;对产品评论在产品特征层面进行了情感分类,并将带有情感分类标签的产品评论进行基于产品特征的聚类,有利于对产品特征情感倾向的分析,该模型为产品评论细粒度情感分析提供了一种有效思路。本文对产品评论采取的是情感二分类方法,对情感倾向进行了积极与消极的判断,未来可深入研究如何实现情感多分类,实现对情感强度的判断。

参考文献:

[1] 王仁武,宋家怡,陈川宝. 基于 Word2Vec 的情感分析在品牌认知中的应用研究[J]. 图书情报工作,2017, 61(22):6－12.

[2] VO D T,ZHANG Y. Don't count,predict! an automatic approach to learning sentiment lexious for short text[C] // YANNICK V, HAI Z, YUSUKE M. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. Stroundsburg: ACL, 2016:219－224.

[3] TANG D Y,QIN B, ZHOU L J,et al. Domain-specific sentiment word extraction by seed expansion and pattern generation[J/OL]. [2019－11－26]. http://arxiv. org/pdf/1309. 6722. pdf.

[4] ZHENG L,WANG H,GAO S. Sentimental feature selection for sentiment analysis of Chinese online reviews[J]. International journal of machine learning and cybernetics,2018,9(1):75－84.

[5] TITOV I,MCDONALD R. Modeling online reviews with multigrain topic models[C]//LI Z. Proceeding of the 17th international conference on World Wide Web. New York: ACM,2008: 111－120.

[6] 孙艳,周学广,付伟. 基于主题情感混合模型的无监督文本情感分析[J]. 北京大学学报(自然科学版),2013,49(1):102－108.

[7] KIM Y. Convolutional neural networks for sentence classification [C]// YUVAL M. Proceedings of the 2014 conference on empirical methods in natural language processing . Doha: ACL, 2014: 1746－1751.

[8] 李杰,李欢. 基于深度学习的短文本评论产品特征提取及情感分类研究[J]. 情报理论与实践,2018(2):143－148.

[9] 余传明,陈雷,张小青. 基于支持向量机的产品属性识别研究[J]. 情报学报,2010,29(6):1038－1044.

[10] 徐建民,王金花,马伟瑜. 利用本体关联度改进的 TF-IDF 特征词提取方法[J]. 情报科学,2011,29(2) :279－283.

[11] 夏天. 词向量聚类加权 TextRank 的关键词提取[J] . 数据分析与知识发现,2017(2) :28－33.

[12] 张柳,王昕巍,黄博,等. 基于字词向量的多尺度卷积神经网络微博评论的情感分类模型及实验研究[J]. 图书情报工作,2019,63(18):99－108.

[13] 李跃鹏,金翠,及俊川. 基于 Word2Vec 的关键字提取算法[J]. 科研信息化技术与应用,2015,6(4):54－59.

[14] 周顺先,蒋励,林霜巧,等. 基于 Word2Vector 的文本特征化表示方法[J]. 重庆邮电大学学报(自然科学版),2018,30(2):272－279.

[15] 金志刚,胡博宏,张瑞. 基于深度学习的多维特征微博情感分析[J]. 中南大学学报(自然科学版). 2018(5):1135－1140.

[16] 赖文辉,乔宇鹏. 基于词向量和卷积神经网络的垃圾短信识别方法[J] . 计算机应用,2018,38(9):2469－2476.

[17] MIKOLOV T,CHEN K,CORRADO G,et al. Efficient estimation of word representations in vector space [J/OL]. [2019－11－26]. http://arxiv. org/pdf/1301. 3781v3. pdf.

[18] MIKOLOV T,SUTSKEVER L,CHEN K,et al. Distributed representations of words and phrases and their compositionality [C] // BURGES C J C, BOTTOU L, WELLING M. Proceedings of the 26th international conference on neural information processing systems. Lake Tahoe: NIPS,2013:3111－3119.

[19] 宁建飞,刘降真. 融合 Word2Vec 与 TextRank 的关键词抽取研究[J]. 数据分析与知识发现,2016(6):20－27.

[20] 张剑,屈丹,李真. 基于词向量特征的循环神经网络语言模型[J]. 模式识别与人工智能,2015(4):299－305.

[21] WU Y B,ZHANGg Q,HUANG X J,et al. Phrase dependency parsing for opinion mining[C] //DAVID C. Proceedings of the 2009 conference on empirical methods in natural language processing. Singapore: ACL,2009(3):1533－1541.

[22] 李志宇,梁循,周小平. 基于属性主题分割的评论短文本词向量构建优化算法[J]. 中文信息学报,2016(9):101－110.

[23] 刘奇飞,沈炜域. 基于 Word2Vec 和 TextRank 的时政类新闻关键词抽取方法研究[J]. 情报探索,2018(6):22－27.

[24] 徐晨,曹辉,赵晓. 基于 SVM 的说话人识别参数选择方法[J]. 计算机工程,2012(11):175－177.

作者贡献说明:

蔡庆平:数据采集与处理、实验、论文初稿撰写与修改;
马海群:论文选题与框架制定,论文修改和审阅。

A Fine-grained Sentiment Analysis Model for Product Reviews Based on Word2Vec and CNN

Cai Qingping[1] Ma Haiqun[2]
[1] School of Information Management, Heilongjiang University, Harbin 150080
[2] Research Center of Information Resource Management, Heilongjiang University, Harbin 150080

Abstract: [Purpose/significance] To construct a fine-grained sentiment analysis model for product reviews based on Word2Vec and CNN. [Method/process] This paper firstly applied Word2vec to build product feature vocabulary and noise vocabulary based on product reviews, secondly extracted the feature words from product reviews by the noise vocabulary, then classified the product reviews according to product features sentiment, finally realized product reviews clustering based on product features. [Result/conclusion] The model was trained and tested by the reviews of Huawei mobile phone on JingDong Mall, the results showed that the model could effectively realize fine-grained sentiment analysis of product reviews and find out users focus and satisfaction on product features.

Keywords: sentiment analysis product reviews CNN Word2Vec fine-grained